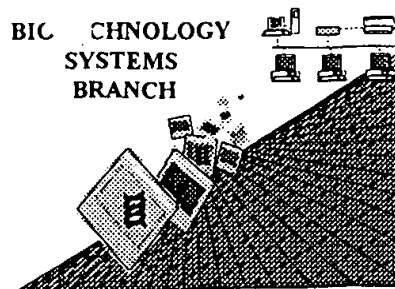


**BEST AVAILABLE COPY**

**RAW SEQUENCE LISTING**  
**ERROR REPORT**



The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/831, 426  
Source: Pct 09  
Date Processed by STIC: 5-23-01

**THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.**

**PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:**

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY.**

**FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.**

**FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.**

**PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)**

**TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.0 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:**

**Checker Version 3.0**

The Checker Version 3.0 application is a state-of-the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 - 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.

Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).

Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

**Checker Version 3.0 can be down loaded from the USPTO website at the following address:**

**<http://www.uspto.gov/web/offices/pac/checker>**

PCT09

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/831,426

DATE: 05/23/2001

TIME: 15:44:24

Input Set : A:\2503.APP.txt

Output Set: C:\CRF3\05232001\I831426.raw

Does Not Comply  
Corrected Diskette Needed  
global errors

3 <110> APPLICANT: Hoechst Marion Roussel  
5 <120> TITLE OF INVENTION: HUMAN htFIII GENE AND CODED htFIIIA PROTEIN  
7 <130> FILE REFERENCE: 9823seq  
C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/831,426  
C--> 10 <141> CURRENT FILING DATE: 2001-05-08  
12 <160> NUMBER OF SEQ ID NOS: 10  
14 <170> SOFTWARE: PatentIn Vers. 2.0

## ERRORED SEQUENCES

16 <210> SEQ ID NO: 1  
17 <211> LENGTH: 1273  
E--> 18 <212> TYPE: ADN → Valid responses: DNA, RNA, PRT  
19 <213> ORGANISM: Human  
21 <220> FEATURE:  
22 <221> NAME/KEY: CDS  
23 <222> LOCATION: (176)..(1270)  
E--> 25 <400> SEQUENCE: 0 → Incorrect: <400> Sequence ID No: 1 Correct: <400> 1  
26 atgcgcagca gcggcgccga cgcgggcggtg cgcctgggtga ccgcgcgcgc tcccgggaagt 60  
28 gtgcccggcgt cgcgcggaagg ttcagcaggg agccgtgggc cgggcgcgcc gggtcccggc 120  
30 acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg gcgcc ctg 178  
31 Leu  
32 1  
34 gat ccg ccg gcc gtg gtc gcc gag tcg gtg tcg tcc ttg acc atc gcc 226  
35 Asp Pro Pro Ala Val Val Ala Glu Ser Val Ser Ser Leu Thr Ile Ala  
36 5 10 15  
38 gac gcg ttc att gca gcc ggc gag agc tca gct ccg acc ccg ccg cgc 274  
39 Asp Ala Phe Ile Ala Ala Gly Glu Ser Ser Ala Pro Thr Pro Pro Arg  
40 20 25 30  
42 ccc gcg ctt ccc agg agg ttc atc tgc tcc ttc cct gac tgc agc gcc 322  
43 Pro Ala Leu Pro Arg Arg Phe Ile Cys Ser Phe Pro Asp Cys Ser Ala  
44 35 40 45  
46 aat tac agc aaa gcc tgg aag ctt gac gcg cac ctg tgc aag cac acg 370  
47 Asn Tyr Ser Lys Ala Trp Lys Leu Asp Ala His Leu Cys Lys His Thr  
48 50 55 60 65  
50 ggg gag aga cca ttt gtt tgt gac tat gaa ggg tgt ggc aag gcc ttc 418  
51 Gly Glu Arg Pro Phe Val Cys Asp Tyr Glu Gly Cys Gly Lys Ala Phe  
52 70 75 80  
54 atc agg gac tac cat ctg agc cgc cac att ctg act cac aca gga gaa 466  
55 Ile Arg Asp Tyr His Leu Ser Arg His Ile Leu Thr His Thr Gly Glu  
56 85 90 95  
58 aag ccg ttt gtt tgt gca gcc act ggc tgt gat caa aaa ttc aac aca 514  
59 Lys Pro Phe Val Cys Ala Ala Thr Gly Cys Asp Gln Lys Phe Asn Thr  
60 100 105 110  
62 aaa tca aac ttg aag aaa cat ttt gaa cgc aaa cat gaa aat caa caa 562  
63 Lys Ser Asn Leu Lys Lys His Phe Glu Arg Lys His Glu Asn Gln Gln

Respond with  
number only,  
no text.

Note

→ The types of errors shown exist throughout the Sequence Listing. Please check  
file://C:\Crf3\ subsequent sequences for similar errors.

5/23/01

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/831,426

DATE: 05/23/2001

TIME: 15:44:24

Input Set : A:\2503.APP.txt

Output Set: C:\CRF3\05232001\I831426.raw

```

64      115      120      125
66 aaa caa tat ata tgc agt ttt gaa gac tgt aag aag acc ttt aag aaa 610
67 Lys Gln Tyr Ile Cys Ser Phe Glu Asp Cys Lys Lys Thr Phe Lys Lys
68 130      135      140      145
70 cat cag cag ctg aaa atc cat cag tgc cag cat acc aat gaa cct cta 658
71 His Gln Gln Leu Lys Ile His Gln Cys Gln His Thr Asn Glu Pro Leu
72      150      155      160
74 ttc aag tgt acc cag gaa gga tgt ggg aaa cac ttt gca tca ccc agc 706
75 Phe Lys Cys Thr Gln Glu Gly Cys Gly Lys His Phe Ala Ser Pro Ser
76      165      170      175
78 aag ctg aaa cga cat gcc aag gcc cac gag ggc tat gta tgt caa aaa 754
79 Lys Leu Lys Arg His Ala Lys Ala His Glu Gly Tyr Val Cys Gln Lys
80      180      185      190
82 gga tgt tcc ttt gtg gca aaa aca tgg acg gaa ctt ctg aaa cat gtg 802
83 Gly Cys Ser Phe Val Ala Lys Thr Trp Thr Glu Leu Leu Lys His Val
84      195      200      205
86 aga gaa acc cat aaa gag gaa ata cta tgt gaa gta tgc cgg aaa aca 850
87 Arg Glu Thr His Lys Glu Glu Ile Leu Cys Glu Val Cys Arg Lys Thr
88 210      215      220      225
90 ttt aaa cgc aaa gat tac ctt aag caa cac atg aaa act cat gcc cca 898
91 Phe Lys Arg Lys Asp Tyr Leu Lys Gln His Met Lys Thr His Ala Pro
92      230      235      240
94 gaa agg gat gta tgt cgc tgt cca aga gaa ggc tgt gga aga acc tat 946
95 Glu Arg Asp Val Cys Arg Cys Pro Arg Glu Gly Cys Gly Arg Thr Tyr
96      245      250      255
98 act act gtg ttt aat ctc caa agc cat atc ctc tcc ttc cat gag gaa 994
99 Thr Thr Val Phe Asn Leu Gln Ser His Ile Leu Ser Phe His Glu Glu
100      260      265      270
102 agc cgc cct ttt gtg tgt gaa cat gct ggc tgt ggc aaa aca ttt gca 1042
103 Ser Arg Pro Phe Val Cys Glu His Ala Gly Cys Gly Lys Thr Phe Ala
104      275      280      285
106 atg aaa caa agt ctc act agg cat gct gtt gta cat gat cct gac aag 1090
107 Met Lys Gln Ser Leu Thr Arg His Ala Val Val His Asp Pro Asp Lys
108 290      295      300      305
110 aag aaa atg aag ctc aaa gtc aaa aaa tct cgt gaa aaa cgg agt ttg 1138
111 Lys Lys Met Lys Leu Lys Val Lys Lys Ser Arg Glu Lys Arg Ser Leu
112      310      315      320
114 gcc tct cat ctc agt gga tat atc cct ccc aaa agg aaa caa ggg caa 1186
115 Ala Ser His Leu Ser Gly Tyr Ile Pro Pro Lys Arg Lys Gln Gly Gln
116      325      330      335
118 ggc tta tct ttg tgt caa aac gga gag tca ccc aac tgt gtg gaa gac 1234
119 Gly Leu Ser Leu Cys Gln Asn Gly Glu Ser Pro Asn Cys Val Glu Asp
120      340      345      350
122 aag atg ctc tcg aca gtt gca gta ctt acc ctt ggc taa 1273
123 Lys Met Leu Ser Thr Val Ala Val Leu Thr Leu Gly
124      355      360      365
127 <210> SEQ ID NO: 2
128 <211> LENGTH: 365
129 <212> TYPE: PRT

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/831,426

DATE: 05/23/2001

TIME: 15:44:24

Input Set : A:\2503.APP.txt

Output Set: C:\CRF3\05232001\I831426.raw

130 <213> ORGANISM: Human  
 E--> 132 <400> SEQUENCE: 0 *See p. 1*

133	Leu	Asp	Pro	Pro	Ala	Val	Val	Ala	Glu	Ser	Val	Ser	Ser	Leu	Thr	Ile
134	1				5				10					15		
136	Ala	Asp	Ala	Phe	Ile	Ala	Ala	Gly	Glu	Ser	Ser	Ala	Pro	Thr	Pro	Pro
137				20					25					30		
139	Arg	Pro	Ala	Leu	Pro	Arg	Arg	Phe	Ile	Cys	Ser	Phe	Pro	Asp	Cys	Ser
140			35					40					45			
142	Ala	Asn	Tyr	Ser	Lys	Ala	Trp	Lys	Leu	Asp	Ala	His	Leu	Cys	Lys	His
143		50					55					60				
145	Thr	Gly	Glu	Arg	Pro	Phe	Val	Cys	Asp	Tyr	Glu	Gly	Cys	Gly	Lys	Ala
146	65					70				75					80	
148	Phe	Ile	Arg	Asp	Tyr	His	Leu	Ser	Arg	His	Ile	Leu	Thr	His	Thr	Gly
149				85					90					95		
151	Glu	Lys	Pro	Phe	Val	Cys	Ala	Ala	Thr	Gly	Cys	Asp	Gln	Lys	Phe	Asn
152				100					105					110		
154	Thr	Lys	Ser	Asn	Leu	Lys	Lys	His	Phe	Glu	Arg	Lys	His	Glu	Asn	Gln
155			115					120					125			
157	Gln	Lys	Gln	Tyr	Ile	Cys	Ser	Phe	Glu	Asp	Cys	Lys	Lys	Thr	Phe	Lys
158		130					135					140				
160	Lys	His	Gln	Gln	Leu	Lys	Ile	His	Gln	Cys	Gln	His	Thr	Asn	Glu	Pro
161	145					150				155					160	
163	Leu	Phe	Lys	Cys	Thr	Gln	Glu	Gly	Cys	Gly	Lys	His	Phe	Ala	Ser	Pro
164				165					170					175		
166	Ser	Lys	Leu	Lys	Arg	His	Ala	Lys	Ala	His	Glu	Gly	Tyr	Val	Cys	Gln
167				180					185					190		
169	Lys	Gly	Cys	Ser	Phe	Val	Ala	Lys	Thr	Trp	Thr	Glu	Leu	Leu	Lys	His
170			195					200					205			
172	Val	Arg	Glu	Thr	His	Lys	Glu	Glu	Ile	Leu	Cys	Glu	Val	Cys	Arg	Lys
173		210					215					220				
175	Thr	Phe	Lys	Arg	Lys	Asp	Tyr	Leu	Lys	Gln	His	Met	Lys	Thr	His	Ala
176	225					230				235				240		
178	Pro	Glu	Arg	Asp	Val	Cys	Arg	Cys	Pro	Arg	Glu	Gly	Cys	Gly	Arg	Thr
179				245					250					255		
181	Tyr	Thr	Thr	Val	Phe	Asn	Leu	Gln	Ser	His	Ile	Leu	Ser	Phe	His	Glu
182				260					265					270		
184	Glu	Ser	Arg	Pro	Phe	Val	Cys	Glu	His	Ala	Gly	Cys	Gly	Lys	Thr	Phe
185			275					280				285				
187	Ala	Met	Lys	Gln	Ser	Leu	Thr	Arg	His	Ala	Val	Val	His	Asp	Pro	Asp
188		290					295					300				
190	Lys	Lys	Lys	Met	Lys	Leu	Lys	Val	Lys	Lys	Ser	Arg	Glu	Lys	Arg	Ser
191	305					310				315					320	
193	Leu	Ala	Ser	His	Leu	Ser	Gly	Tyr	Ile	Pro	Pro	Lys	Arg	Lys	Gln	Gly
194				325					330					335		
196	Gln	Gly	Leu	Ser	Leu	Cys	Gln	Asn	Gly	Glu	Ser	Pro	Asn	Cys	Val	Glu
197				340					345					350		
199	Asp	Lys	Met	Leu	Ser	Thr	Val	Ala	Val	Leu	Thr	Leu	Gly			
200			355					360					365			
203	<210>	SEQ ID NO:	3													

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/831,426

DATE: 05/23/2001

TIME: 15:44:24

Input Set : A:\2503.APP.txt

Output Set: C:\CRF3\05232001\I831426.raw

204 <211> LENGTH: 1273

E--> 205 <212> TYPE: ADN *See p. 1*

206 <213> ORGANISM: Human

E--> 208 <400> SEQUENCE: 0

```

209 atgcgcagca gcggcgccga cgcggggcgg tgccctggtga ccgcgcgcgc tcccggaaagt 60
211 gtgccggcgt cgcgcgaaagg ttcagcaggg agccgtgggc cgggcgcgcc ggttcccggc 120
213 acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg gcgccctgga 180
215 tccgccggcc gtggtcgccg agtcggtgtc gtccttgacc atcgccgacg cgttcattgc 240
217 agccggcgag agctcagctc cgaccccgcc gcgccccgcg cttcccagga ggttcatctg 300
219 ctccctccct gactgcagcg ccaattacag caaagcctgg aagcttgacg cgcacctgtg 360
221 caagcacacg ggggagagac catttgtttg tgactatgaa ggggtgtggca aggccttcat 420
223 cagggactac catctgagcc gccacattct gactcacaca ggagaaaagc cgtttgtttg 480
225 tgcagccact ggctgtgatc aaaaattcaa cacaaaatca aacttgaaga aacattttga 540
227 acgcaaacat gaaaatcaac aaaaacaata tatatgcagt tttgaagact gtaagaagac 600
229 ctttaagaaa catcagcagc tgaaaatcca tcagtgccag cataccaatg aacctctatt 660
231 caagtgtacc caggaaggat gtgggaaaca ctttgcatca cccagcaagc tgaaacgaca 720
233 tgccaaggcc cacgagggct atgtatgtca aaaaggatgt tcctttgttg caaaaacatg 780
235 gacggaactt ctgaaacatg tgagagaaac ccataaagag gaaatactat gtgaagtatg 840
237 ccgaaaaaca tttaaacgca aagattacct taagcaacac atgaaaactc atgccccaga 900
239 aagggatgta tgtcgctgtc caagagaagg ctgtggaaga acctatacta ctgtgtttaa 960
241 tctccaaagc catatcctct ccttccatga ggaaagccgc cttttgtgtg gtgaacatgc 1020
243 tggctgtggc aaaacatttg caatgaaaca aagtctcact aggcattgctg ttgtacatga 1080
245 tcctgacaag aagaaaatga agctcaaagt caaaaaatc cgtgaaaaac ggagtttggc 1140
247 ctctcatctc agtggatata tccctcccaa aaggaaacaa gggcaaggct tatctttgtg 1200
249 tcaaaacgga gagtaccca actgtgtgga agacaagatg ctctcgacag ttgcagtact 1260
251 tacccttggc taa
254 <210> SEQ ID NO: 4
255 <211> LENGTH: 1213
E--> 256 <212> TYPE: ADN See p. 1
257 <213> ORGANISM: Human
E--> 259 <400> SEQUENCE: 0
260 gtgccggcgc cgcgcgaaagg ttcagcaggg agccgtgggc cgggcgcgcc ggttcccggc 60
262 acgtgtctcg gcacgtggca gcgcgcctgg ccctgggctt ggaggcgccg gcgccctgga 120
264 tccgccggcc gtggtcgccg agtcggtgtc gtccttgacc atcgccgacg cgttcattgc 180
266 agccggcgag agctcagctc cgaccccgcc gcgccccgcg cttcccagga ggttcatctg 240
268 ctccctccct gactgcagcg ccaattacag caaagcctgg aagcttgacg cgcacctgtg 300
270 caagcacacg ggggagagac catttgtttg tgactatgaa ggggtgtggca aggccttcat 360
272 cagggactac catctgagcc gccacattct gactcacaca ggagaaaagc cgtttgtttg 420
274 tgcagccact ggctgtgatc aaaaattcaa cacaaaatca aacttgaaga aacattttga 480
276 acgcaaacat gaaaatcaac aaaaacaata tatatgcagt tttgaagact gtaagaagac 540
278 ctttaagaaa catcagcagc tgaaaatcca tcagtgccag cataccaatg aacctctatt 600
280 caagtgtacc caggaaggat gtgggaaaca ctttgcatca cccagcaagc tgaaacgaca 660
282 tgccaaggcc cacgagggct atgtatgtca aaaaggatgt tcctttgttg caaaaacatg 720
284 gacggaactt ctgaaacatg tgagagaaac ccataaagag gaaatactat gtgaagtatg 780
286 ccgaaaaaca tttaaacgca aagattacct taagcaacac atgaaaactc atgccccaga 840
288 aagggatgta tgtcgctgtc caagagaagg ctgtggaaga acctatacta ctgtgtttaa 900
290 tctccaaagc catatcctct ccttccatga ggaaagccgc cttttgtgtg gtgaacatgc 960
292 tggctgtggc aaaacatttg caatgaaaca aagtctcact aggcattgctg ttgtacatga 1020
294 tcctgacaag aagaaaatga agctcaaagt caaaaaatc cgtgaaaaac ggagtttggc 1080

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/831,426

DATE: 05/23/2001

TIME: 15:44:24

Input Set : A:\2503.APP.txt

Output Set: C:\CRF3\05232001\I831426.raw

```

296 ctctcatctc agtggatata tcctcccaa aaggaaacaa gggcaaggct tatctttgtg 1140
298 tcaaaacgga gagtcacca actgtgtgga agacaagatg ctctcgacag ttgcagtact 1200
300 tacccttggc taa 1213
303 <210> SEQ ID NO: 5
304 <211> LENGTH: 34
E--> 305 <212> TYPE: ADN See p.1
306 <213> ORGANISM: Human
E--> 308 <400> SEQUENCE: 0
309 cggggtacca aaaatgcgca gcagcggcgc cgac 34
312 <210> SEQ ID NO: 6
313 <211> LENGTH: 21
E--> 314 <212> TYPE: ADN See p.1
315 <213> ORGANISM: Human
E--> 317 <400> SEQUENCE: 0
318 tccttcctg actgcagcgc c 21
321 <210> SEQ ID NO: 7
322 <211> LENGTH: 20
E--> 323 <212> TYPE: ADN See p.1
324 <213> ORGANISM: Human
E--> 326 <400> SEQUENCE: 0
327 tgcacagggtg cgcgtcaagc 20
330 <210> SEQ ID NO: 8
331 <211> LENGTH: 20
E--> 332 <212> TYPE: ADN See p.1
333 <213> ORGANISM: Human
E--> 335 <400> SEQUENCE: 0
336 cacaaacaaa tggctctctc 20
339 <210> SEQ ID NO: 9
340 <211> LENGTH: 30
E--> 341 <212> TYPE: ADN See p.1
342 <213> ORGANISM: Human
E--> 344 <400> SEQUENCE: 0
345 cggctagat tagccaagg taagtactgc 30
348 <210> SEQ ID NO: 10
349 <211> LENGTH: 30
E--> 350 <212> TYPE: ADN See p.1
351 <213> ORGANISM: Human
E--> 353 <400> SEQUENCE: 0
354 cctcccgggg ccaagggtaa gtactgcaac 30

```

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/831,426

DATE: 05/23/2001

TIME: 15:44:25

Input Set : A:\2503.APP.txt

Output Set: C:\CRF3\05232001\I831426.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number  
L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date  
L:18 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:25 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:1 differs:0  
L:132 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:2 differs:0  
L:205 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:208 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:3 differs:0  
L:256 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:259 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:4 differs:0  
L:305 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:308 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:5 differs:0  
L:314 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:317 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:6 differs:0  
L:323 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:326 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:7 differs:0  
L:332 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:335 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:8 differs:0  
L:341 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:344 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:9 differs:0  
L:350 M:310 E: (3) Wrong or Missing Sequence Type, TYPE:  
L:353 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:10 differs:0